AD_____
(Leave blank)

Award Number:
W81XWH-08-1-0614

TITLE: Understanding the Biosynthesis of SF2575: A Potent Antitumor Compound With Novel Modes of Action

PRINCIPAL INVESTIGATOR:
Yi Tang, Ph.D.

CONTRACTING ORGANIZATION:
University of California Los Angeles
Los Angeles, CA 90095

REPORT DATE:
September 2009

TYPE OF REPORT:
Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: (Check one)

X Approved for public release; distribution unlimited

☐ Distribution limited to U.S. Government agencies only; report contains proprietary information

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
09-01-09	Final	1 SEP 2008-31 AUG 2009

4. TITLE AND SUBTITLE Understanding the Biosynthesis SF2575: A Potent Antitumor Compound with Novel Modes of Action

5a. CONTRACT NUMBER W81XWH-08-1-0614

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Yi Tang
Email: yitang@ucla.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) AND ADDRESS(ES)

*
Univ. of Calif. Los Angeles
11000 Kinross Avenue, Suite 102
Los Angeles, CA 90095

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
DEPARTMENT OF THE ARMY
US ARMY MEDICAL RESEARCH AND MATERIEL COMMAND
504 SCOTT STREET
FORT DETRICK, MD 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Unlimited Distribution

13. SUPPLEMENTARY NOTES

14. ABSTRACT: SF2575 is a tetracycline polyketide produced by *Streptomyces* sp. SF2575 and displays exceptionally potent anticancer activity. The structure is chemically complex and contains unusual angelate and salicylate tailoring groups. In this study, we identified, sequenced and functionally analyzed the *ssf* biosynthetic gene cluster. Intermediates were isolated from the SF2575 culture extract to suggest the order of pendant groups addition is C-9 glycosylation, C-4 salicylation and C-4' angelycylation. Cytotoxicity studies demonstrated loss of activity following removal of angelate and further inactivity following loss of salicylate indicating that both acyl groups are critical to the potent bioactivity. Two enzymes responsible for C-4 acylation of salicylate were identified and characterized in vitro, an ATP-dependent salicylyl-CoA ligase SsfL1 and acyltransferase SsfX3. Understanding the biosynthesis of SF2575 can therefore expand the repertoire of enzymes that can modify tetracyclines, and facilitate engineered biosynthesis of SF2575 analogs.

15. SUBJECT TERMS
Anticancer, chemotherapy, Natural product, topoisomerase

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	17	19b. TELEPHONE NUMBER *(include area code)*

Table of Contents

Introduction

Natural products produced by bacteria and fungi encompass a broad range of bioactivity and are an important source of therapeutics in use today. Cancer therapy is no exception to this. One prominent example is doxorubicin, a compound naturally produced by a soil-dwelling bacterium, and is a common component in chemotherapy regimens for many types of cancers including breast cancer. Due to the chemical complexity of these molecules which often prohibits de novo chemical synthesis, biosynthesis is an attractive approach to produce and generate analogs of these bioactive compounds. SF2575, studied here, is a tetracycline family natural product which was found to have exceptionally potent anticancer activity both through an in vitro cytotoxicity assay with P388 leukemia cells, and during an in vivo assay using mouse xenografts[1]. More recently, a 60-cell line screening by the National Cancer Institute demonstrated the potent activity of SF2575 against nearly all types of cancer cell lines tested, resulting in an average IC_{50} value of 11.2 nM. The mechanism of action of several closely related SF2575 analogs, TAN-1518A and TAN-1518B has been identified as inhibition of DNA topoisomerase I[2]. As these structures are nearly identical to that of SF2575, it is likely that this family of compounds shares a common molecular target. As key enzymes during DNA replication, both DNA topoisomerase I and II are known targets for current anticancer therapies in clinical use such as doxorubicin, a topoisomerase II poison[3], and camptothecin derivatives which target topoisomerase I[4].

During this study, we have identified and sequenced the *ssf* gene cluster responsible for the biosynthesis of SF2575 from *Streptomyces* sp. SF2575. This *ssf* gene cluster is only the third tetracycline family gene cluster to be identified and sequenced. The genetic information offers valuable opportunities to further enhance our understanding of tetracycline biosynthesis, and to investigate an entirely new set of tetracycline tailoring modifications that distinguish SF2575 from the previously studied oxytetracycline and chlorotetracycline (Figure 1). SF2575 has a much more complex and heavily decorated structure than previously studied tetracyclines which is reflected in the size of the sequenced gene cluster which was nearly double that of oxytetracycline. Bioinformatic analysis of the gene cluster along with identification of the possible biosynthetic intermediates from *S.* sp. SF2575 fermentation extract led to the assembly of a putative biosynthetic pathway for synthesis of SF2575. We have reconstituted the early portions of the *ssf* pathway in a heterologous host and demonstrated the biosynthesis of the tetracycline core parallels the previously published biosynthetic pathway of oxytetracycline[5-7]. Additionally, we have identified key tailoring enzymes involved in the attachment of salicylate to the tetracycline aglycon and verified their function through in vitro analysis. Elucidation of the SF2575 biosynthetic pathway and characterization of these key tailoring enzymes is an important first step toward structure activity relationship studies and engineered biosynthesis of new anticancer compounds.

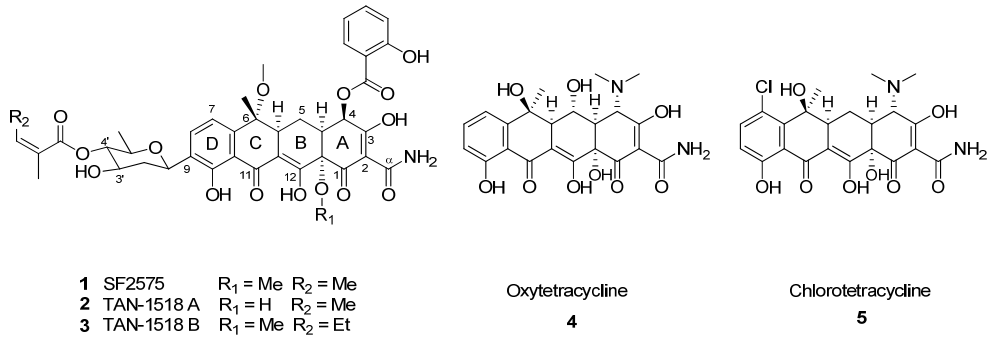

Figure 1: Natural product tetracyclines: SF2575 **1** produced by *S. sp.* SF2575; TAN-1518A **2** and TAN-1518B **3** produced *by Streptomyces sp.* AL-16012; oxytetracycline **4** produced by *S. rimosus*; and chlorotetracycline **5** produced by *S. aureofaciens.*

Body

Task 1. Elucidate the biosynthetic pathway of SF2575 from the native producer.

Identification and sequencing of gene cluster

Aromatic polyketide biosynthetic genes are typically clustered together on the host genomes, which makes them ideal for identification by PCR screening of cosmid libraries[8]. Degenerate primers based on the conserved KS_α sequence[9] were used to probe a cosmid library containing *S.* sp. SF2575 genomic DNA. Several PCR products from different cosmids were sequenced and were found to encode an identical KS fragment. This indicates that there is likely only one aromatic PKS gene cluster present in the *S.* sp. SF2575 genome.

Cosmid 5F-15 was selected and sequenced using a combination of shotgun sequencing and primer walking. Sequences near the ends of this genomic fragment were used to screen for additional cosmids containing overlapping sequence information. A double crossover mutant which inactivated *ssfB* was created to verify the authenticity of the gene cluster. The resulting mutant produced no SF2575 or any recognizable SF2575 precursors, which indicates that the *ssf* gene cluster is indeed responsible for SF2575 biosynthesis. The *ssf* biosynthetic cluster spans a 47.2 kb region and contains 40 ORFs (Figure 2). Functions of the encoded proteins were assigned by sequence similarity using NCBI protein-protein BLAST (Table A1). A putative biosynthetic pathway was then constructed using the proposed protein functions. Heterologous biosynthesis and in vitro assays were used to confirm sections of the proposed pathway including biosynthesis of the tetracycline core and attachment of the salicylic acid ligand. The KS II gene was knocked out by a double crossover mutation and the absence of SF2575 production from the resulting culture confirmed that this gene cluster is indeed responsible for SF2575 biosynthesis.

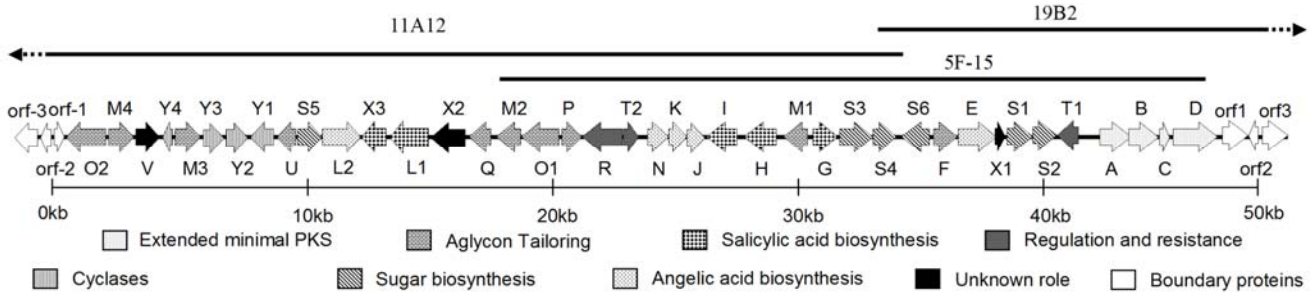

Figure 2: Organization of the *ssf* biosynthetic gene cluster. Genes are categorized according to their proposed role. The gene cluster spans 47.2 kb and contains 40 ORFs. Details of proposed functions are shown in Table 1.

Identification of putative intermediates from fermentation broth

In order to decipher the timing of the three tailoring reactions in the biosynthesis of SF2575: glycosylation with D-olivose, acylation of salicylic acid and acylation of angelic acid, crude extract of *S.* sp. SF2575 culture was prepared and analyzed on HPLC and LCMS to try to identify any potential stable intermediates that may be present (Figure 3). Using selected ion monitoring, the only potential intermediates identified were one with molecular weight of 575 ($[M+H]^+$ at *m/z* 576, RT = 16.8 min) and a second with molecular weight of 695 ($[M+H]^+$ at *m/z* 696, RT = 23.2 min), in addition to the parent compound SF2575 ($[M+H]^+$ at *m/z* 778, RT = 30.1 min) (Figure 3). The UV spectrum of each putative precursor compound was similar to that of SF2575 with a characteristic tetracycline λ_{max} at 358

nm, and a λ_{max} at 302 nm for SF2575 and **34** characteristic of the salicylate. The UV spectrum of **33** lacks this contribution at 302 nm resulting in a smooth peak at 358 nm, and is indicative of the loss of salicylate compared to **34** as suggested by the molecular weight difference. To confirm the identity of these compounds as shown in Figure 6, authentic standards were prepared from base hydrolysis of purified SF2575 as described by Hatsu et. al.[10] Following purification of 33 and 34 from the base hydrolysis reactions, proton NMR and HRMS were used to confirm the structures by comparison to published[10]. These prepared samples were then used as standards to verify the identities of 33 and 34 in the fermentation extract of *S.* sp. SF2575 by HPLC retention time, mass fragmentation pattern and UV spectra.

Other potential intermediates, such as **33** acylated with angelic acid, or **25** acylated with salicylic acid were not detected in the crude extracts. These evidences, along with the positive detection of **33** and **34** are consistent with the proposed biosynthetic pathway of SF2575 as shown in Figure 6. Therefore, the likely order of the tailoring modifications starts with *C*-glycosylation of **25** with **30** that results in **33**, followed by acylation of **31** at C-4 to produce **34**, and capped by the acylation of O-4' of **34** with **35** to produce the final product SF2575. It is noteworthy that each of these isolated intermediates has already been *O*-methylated at the C-12a and C-6 hydroxyl groups, which indicates that these modifications likely take place early in the biosynthesis, further confirming **25** as the proposed aglycon. While the tetracyclic aglycon was not identified in the fermentation extract, this could be due either to the potential instability of this compound or the rate of the glycosylation reaction which may not permit its accumulation in detectable quantities.

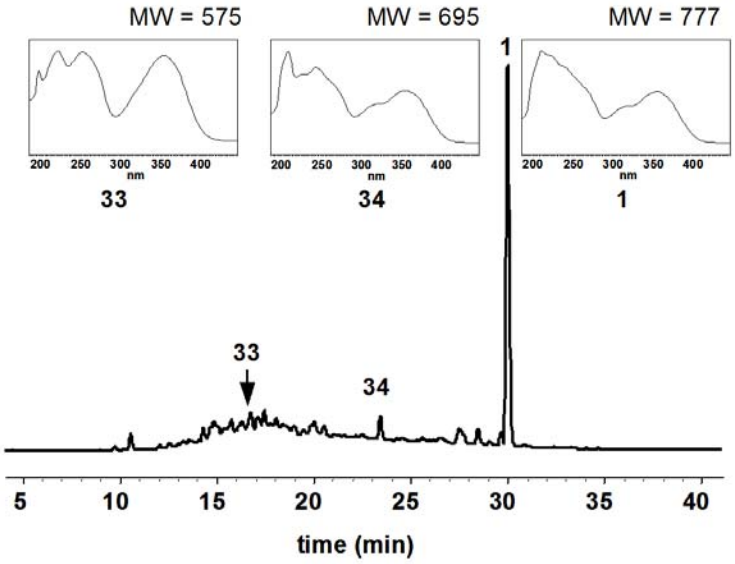

Figure 3 HPLC analysis of the *S.* sp. SF2575 extract after 7 days growth on solid Bennette's media. The main product **1,** along with putative intermediates **33** and **34** are indicated. The UV spectra of these compounds are shown for comparison.

Biosynthesis of the amidated polyketide backbone

Using the information gathered from the fermentation analysis, we constructed a putative pathway for SF2575 using the functionality found in the SF2575 gene cluster shown in figures 4 and 6. The first steps are the biosynthesis of the amidated polyketide backbone. Biosynthesis of SF2575 is predicted to be parallel to that of oxytetracycline and chlorotetracycline through the formation of the tetracycline core. Examination of the carbon backbone of these molecules indicates that the biosynthetic pathway is likely identical through the formation of 4-keto-anhydrotetracycline (4-keto-ATC). Genes putatively involved in the biosynthesis of the SF2575 aglycon share high sequence homology to many of the genes found in the recently elucidated oxytetracycline (*oxy*) biosynthetic pathway[5, 11]. As in the oxytetracycline pathway, amidated polyketide backbone is produced by the minimal polyketide synthase (PKS). The *ssf* minimal PKS consists of SsfA (KS), SsfB (CLF), and SsfC (ACP) and are highly homologous to the *oxy* minimal PKS[5]. The amidotransferase SsfD, which is responsible for producing

the malonamate starter unit unique to tetracyclines, is found adjacent to the *ssf* minimal PKS. The C-2 amidated starter unit is one of the signature moieties of tetracycline, and the revelation of SsfD as an OxyD homolog with 68% sequence identity was one of the earliest convincing pieces of evidence that this gene cluster indeed encodes a tetracycline compound. Zhang and coworkers demonstrated the presence of OxyD was essential for incorporation of the C-2 amide[5]. These four enzymes, SsfABCD, complete the 'extended minimal PKS' sufficient for producing the amidated polyketide chain. In addition, SsfU, which is 76% identical to OxyJ, is predicted to be the C-8 ketoreductase that regioselectively reduces the nascent backbone as shown in figure 4[7].

These function of these genes was verified experimentally through heterologous reconstitution. SsfABCDU were cloned into a pRM5 derived shuttle vector to produce pLP27 and expressed in *S. lividans* K4-114. The product profile of K4-114/pLP27 was indistinguishable from that of K4-114/pWJ35, which contains the corresponding enzymes from the *oxy* biosynthetic pathway[5]. By comparison to an authentic standard, the major product of the extract (RT = 13.5 min) was confirmed to be the isoquinolone WJ35 (20 mg/L)**,** which can form via the spontaneous cyclization of the reduced amidated polyketide **9**. We further investigated whether the *oxy* and *ssf* extended minimal PKS components are functionally interchangeable. Three additional shuttle vectors were prepared with a combination of *oxy* and *ssf* genes and transformed into K4-114. Each of these host/vectors combinations were capable of producing WJ35 in similar quantities as K4-114/pLP27, indicating that the *ssf* components are functionally compatible and equivalent to their *oxy* counterparts. Importantly, these results confirm i) the extended *ssf* minimal PKS (SsfABCD) is capable of synthesizing the full length amidated polyketide precursor; ii) the hypothesis that the carbon backbone of SF2575 is biosynthesized via a tetracycline-like pathway; and iii) SsfU as the ketoreductase.

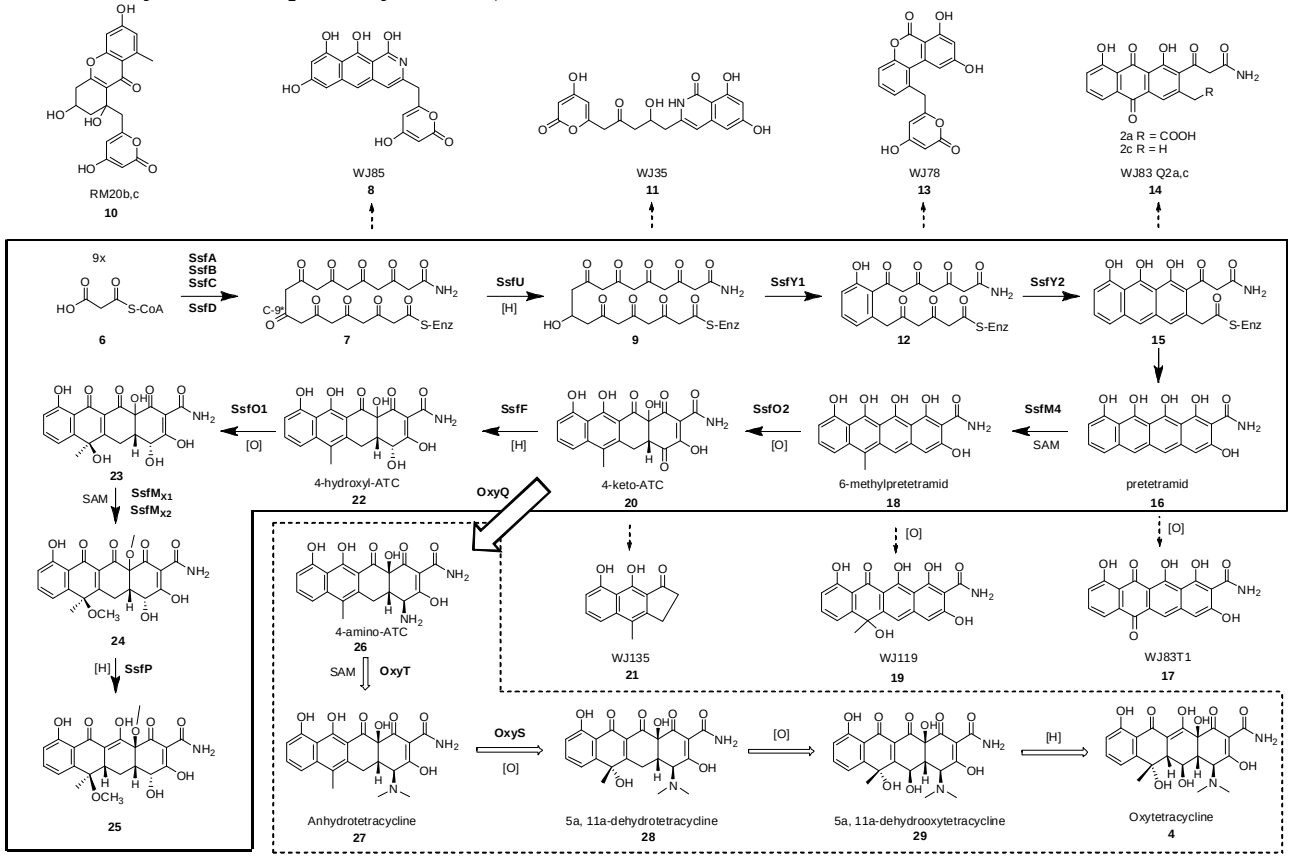

Figure 4. Biosynthesis of the SF2575 aglycon **25**

6-methylpretetramid (6-MPT) is a characteristic stable intermediate identified in tetracycline producing strains. Again, here we expected the formation of 6-MPT during SF2575 biosynthesis to parallel the oxytetracycline pathway. During oxytetracycline biosynthesis, the steps following the extended minimal PKS and reduction by C-8 KR are regioselective cyclization by OxyK and OxyN followed by C-6 methylation to achieve 6-MPT biosynthesis. Unlike the *oxy* cluster, however, the *ssf* gene cluster contains four genes that have homology to cyclases. SsfY1 is highly similar to OxyK (67% sequence similarity) and is predicted to be responsible for D ring cyclization to afford **12**. Indeed, cotransformation of a pSET152 derived integrating vector pLP77 containing SsfY1 with pLP27 into K4-114 resulted in the biosynthesis of **12**, confirming the role of SsfY1. Intriguingly, an additional cyclase, SsfY3, also displays sequence similarity to first ring cyclase/aromatases, with 39% similarity to OxyK. However cotransformation of a SsfY3 containing vector with pLP27 did not result in first ring aromatization, indicating that SsfY3 cannot functionally replace SsfY1 as first ring aromatase/cyclase. Of the remaining two, SsfY2 is homologous to second ring cyclases from several aromatic polyketide pathways. Although SsfY2 has low sequence homology (25% similarity and 14.5% identity) to the second ring cyclase from the *oxy* pathway (OxyN), it displays stronger homology to second ring cyclases from benzoisochromanequinone (BIQ) pathways such as granaticin (68% similarity, 56% identity)[12] and medermycin (66% similarity, 55% identity)[13]. SsfY2 was therefore predicted to be the second ring cyclase in the *ssf* pathway. The last cyclase SsfY4 is similar to OxyI, itself similar to fourth ring cyclase MtmX from the mithramycin pathway, however OxyI was found to be uninvolved in the biosynthesis of oxytetracycline[7].

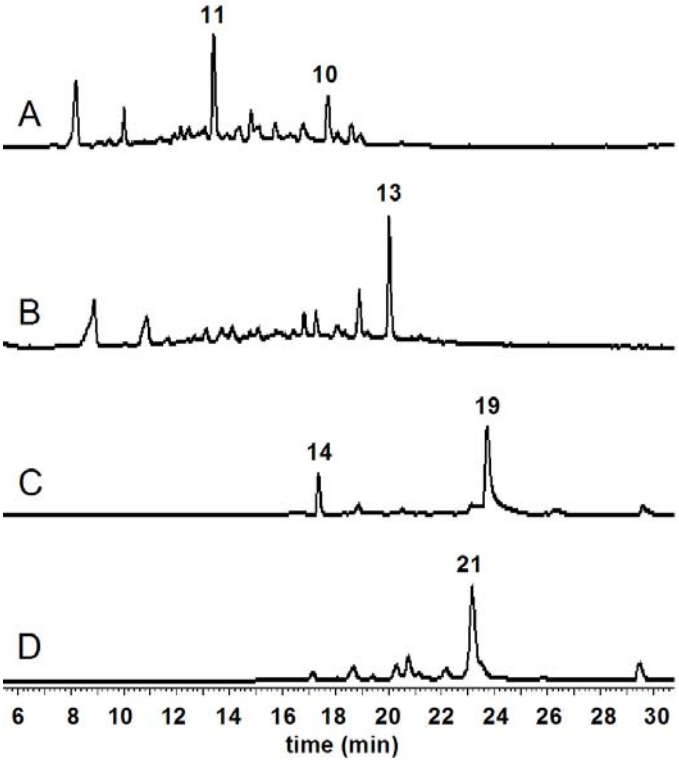

Figure 5: Heterologous reconstitution of early biosynthetic pathway *S. lividans* K4-114. (**A**) HPLC analysis (245 nm) of the K4-114/pLP27 extract shows the amidated, reduced polyketide **11** is the major product, confirming the biosynthesis of the polyketide backbone **7** by SsfABCD and the subsequent C-9* reduction by SsfU. (**B**) HPLC analysis (253 nm) of the K4-114/pLP27/pLP77 extract shows addition of the putative cyclase SsfY1 leads to complete cyclization and aromatization of the D ring and formation of the shunt benzopyrone **13**. (**C**) HPLC analysis (430 nm) of K4-114/pLP27/pLP126 extract shows **19**, the oxidized form of **18,** as the dominant product. Biosynthesis of **18** using entirely *ssf* genes (SsfABCDUY1Y2M4L2) indicates the tetracycline nature of the *ssf* biosynthetic pathway. (**D**) HPLC analysis (395 nm) of K4-114/pLP36 extract confirms the function of SsfO2 as the oxygenase that dihydroxylated C-4 and C-12a

In an attempt to fully reconstitute the biosynthesis of 6-methylpretetramid intermediate, SsfY2 and C-6 methyltransferase SsfM4 were added to pLP77 and cotransformed with pLP27 into K4-114. The resulting strain however produced only three ringed compounds previously identified as WJ83Q2 and WJ83Q1. To determine if the *ssf* pathway requires the additional cyclases found in its gene cluster towards the formation of **16**, pLP115 and pLP118 were constructed, which contained in addition to genes encoded in pLP113, *ssfY4*, and *ssfY3Y4*, respectively. K4-114 contransformed with pLP27 and these constructs continued to produce **14** only, suggesting that for the

ssf pathway, different enzyme(s) may be needed to facilitate cyclization of the fourth ring. We tested this possibility by inserting additional *ssf* genes juxtaposed to the *ssf* cyclase genes in the cluster into K4-114 and examined the metabolic profiles. Fortuitously, upon insertion of *ssfL2* into pLP113 to afford pLP126, and cotransformation of pLP27 and pLP126 into K4-114, the formation of **19** was detected as shown in Figure 5D. SsfL2 has homology to long-chain fatty acyl-CoA ligases as well as putative acyl-CoA ligases from several aromatic polyketide pathways such as CmmLII from the chromomycin gene cluster[14], PokL from the polyketomycin gene cluster[15], MtmL from the mithramycin gene cluster[16], as well as OxyH[5]. The role of these enzymes in their respective biosynthetic pathways has not been determined and has been suggested to function as an accessory protein[14]. Interestingly, biosynthesis of these compounds all require the equivalent Claisen cyclization of the A rings. The mechanism of SsfL2 in the heterologous biosynthesis of **18** with *ssf* PKS components is yet to be determined.

Tailoring steps leading to the putative SF2575 aglycon

During oxytetracycline biosynthesis, the next step following the formation of 6-MPT is the dual hydroxylation of the A ring by OxyL at C-12a and C-4 to produce the cyclohexenone containing 4-keto-anhydrotetracycline (4-keto-ATC) **20**, which is unstable and degrades into observed product WJ135 **21**[6]. SsfO2, which is found at the boundary of the *ssf* pathway, is an OxyL homolog and is most likely to perform this role in the *ssf* pathway. 4-keto-ATC is likely the last common intermediate between the *oxy* and the *ssf* biosynthetic pathways. In the *oxy* pathway, reductive amination of the C-4 ketone group by OxyQ to yield 4-amino-ATC **26** and subsequent dimethylation by OxyT afford the stable intermediate anyhydrotetracycline **27**[6]. The C-4 position in **1** lacks the dimethylamino group and is instead acylated with a salicylic acid group. This requires ketoreduction of **20** at C-4 to produce 4-hydroxyl-ATC **22** for the downstream salicylation reaction. To determine whether oxygenase OxyL installs the hydroxyl group directly or if an additional ketoreductase is necessary, *oxyL* gene was replaced with *ssfO2* in the shuttle plasmid pWJ135, which contains all the *oxy* genes required to synthesize **20** (Table 2)[6]. K4-114 transformed with the resulting plasmid pLP36 produced the tricyclic ketone **21** (Figure 5D), which is a spontaneously degraded product of **20** as observed previously[6]. Then, the genes *oxyQ* and *oxyT* were introduced to pLP36 to afford pLP75 (Table 2). As expected, K4-114/pLP75 produced ~ 10 mg/L of ATC as the dominant product, demonstrating that OxyL and SsfL2 catalyze the identical transformation of **18** to **20.** These results also support the proposal that a dedicated C-4 ketoreductase is required to reduce **20** to **22** and reveal the reactive hydroxyl nucleophile for the subsequent salicylation modification. SsfF is a ketoreductase with sequence homology to NADPH-dependent oxidoreductases with specificity for aromatic and nonpolar substrates (PF00248), and is therefore assigned as a candidate to catalyze this step.

The last tailoring enzyme in the *ssf* pathway that resemblances to enzymes found in *oxy* pathway is SsfO1 which has 54% similarity to OxyS. SsfO1 therefore likely plays the parallel role to oxidize the aromatic C-ring in **22** to **23**. Two *O*-methyltransferases are then needed to methylate the C-6 and the C-12a hydroxyl groups. There are three methyltransferases (other than the C-6 methyltransferase SsfM4) present in the *ssf* gene cluster. SsfM1 and SsfM2 bear strong resemblance to *O*-methyltransferases, while SsfM3 is similar to both *O*-methyltransferases and *C*-methyltransferases. It is likely that two of these are responsible for the *O*-methylation of C-12a and C-6 hydroxyls. The gene encoding SsfM2 is transcribed from the same operon that encodes *ssfO1* and is therefore the likely C-6 *O*-methyltransferase. In the same operon is the gene *ssfQ*, which encodes an S-adenosylmethionine (SAM) synthases and likely serves as an auxiliary role to produce sufficient SAM for the methyltransfer

reactions. It is known that 6-deoxy analogs of tetracyclines, such as doxycycline, have enhanced stability toward acid and base degradation[17]. The C-6-OH methylation may have a similar stabilizing effect on SF2575. Once identified, the C-6 *O*-methyltransferase may potentially be useful to generate C-6-methoxytetracycline analogs that have enhanced stability.

The last step in the biosynthesis of oxytetracycline is the reduction of 5a,11a-dehydrooxytetracycline to oxytetracycline. This reductive modification required for the biosynthesis of chlorotetracycline and oxytetracycline has been postulated to involve a gene located outside their respective gene clusters[18]. The *tchA* gene encodes a flavin dependent oxidoreductase and has been identified in *S. aureofaciens* to be involved for the biosynthesis of **5***[19]*.* An identical reduction step is also needed to complete the biosynthesis of aglycon **25,** from for example, **24**. One candidate for this reaction in the *ssf* gene cluster is a putative dehydrogenase SsfP. Experimental evidence will be needed to determine if this reduction step is catalyzed by an enzyme encoded in the *ssf* gene cluster such as SsfP or if an external gene is required as in chlorotetracycline biosynthesis.

Genes involved in the biosynthesis of D*-olivose*

Deoxysugar decoration is an important tailoring modification of bacterial secondary metabolites and is usually essential for the bioactivity of the natural products[20]. The C-9 of **1** is modified with a *C*-glycoside D-olivose, which is also found among other polyketides such as mithramycin[21], landomycin A[22], urdamycin A[23], and simocyclinone[24] etc. Putative enzymes encoded in the *ssf* gene cluster that can convert glucose to NDP-D-olivose are SsfS1 (NDP-glucose synthase), SsfS2 (4',6'-dehydratase), SsfS3 (2',3'-dehydratase), SsfS4 (C-3' reductase), and SsfS5 (C-4' reductase). NDP-D-olivose is subsequently transferred to the tetracycline aglycon **25** by a *C*-glycosyltransferase to yield **33** (Figure 6). *C*-glycosyltransferases are more rare than *O*-glycosyltransferases, and are found in a number of aromatic PKS biosynthetic gene clusters such as urdamycin[23], hedamycin[25], and simocyclinone[24]. The putative *ssf* glycosyltransferase, SsfS6, has high sequence homology to HedJ, a *C*-glycosyltransferase from the hedamycin biosynthetic pathway[25]. Since SsfS6 is the first glycosyltransferase that can modify the tetracycline scaffold, it can potentially be a useful enzyme towards increasing the structural diversity of the D-ring of tetracycline compounds.

Genes involved in the biosynthesis of angelyl-CoA

Little has been reported on the biosynthesis of angelic acid among bacterial natural products, therefore there is not a well established biosynthetic pathway. Angelate is commonly found to be associated with plant metabolites and is biosynthesized from the isoleucine catabolic pathway[26, 27]. No genes were found in the *ssf* cluster that indicate a plant-like pathway for this unusual substitution in SF2575. Instead, a set of genes encoding enzymes involved in the formation and tailoring of a short chain acyl group was found. These genes include SsfE (carboxytransferase), SsfJ (enoyl-CoA hydratase/isomerase), SsfN (KSIII), and SsfK (3-ketoacyl-ACP reductase). These genes resemble those found in fatty acid biosynthesis and may be responsible for synthesizing angelyl-CoA from a propionyl-CoA, or other short chain acyl building block. One possible pathway to produce angelate from genes found in the *ssf* pathway is shown in figure 6.

SsfE contains a carboxyltransferase domain (PF01039) and is homologous to biotin-dependent methylmalonyl-CoA decarboxylase and propionyl-CoA carboxylase. One possible role for SsfE is to synthesize methylmalonyl-CoA from propionyl-CoA. The next set of steps in the biosynthesis of 35 are proposed to catalyzed by *ssfK, ssfN* and *ssfJ,* which are cotranscribed as a tricistronic cassette on the

same operon. SsfN is homologous to KS III enzymes such as FabH, which catalyzes the condensation between acetyl-CoA and malonyl-ACP to initiate fatty acid biosynthesis[28]. FabH homologs have been found in a number of polyketide gene clusters such as in R1128[29], frenolicin[29], and daunorubicin[30] to catalyze the decarboxylative condensation of short acyl groups. SsfN may therefore catalyze the condensation of methylmalonyl-CoA with acetyl-CoA to form 2-methyl-acetoacetyl-CoA. A β-ketoreduction catalyzed by SsfK, which is homologous to 3-oxoacyl-ACP reductases such as FabG[31], yields 3-hydroxyl-2-methyl-butyryl-CoA**.** Stereospecific dehydration by SsfJ, which is a member of the enoyl-CoA hydratase/isomerase family (PF00378), results in angelate. As an alternative start to this pathway, SsfN may directly condense malonyl-CoA with acetyl-CoA to form acetoacetyl-CoA. The α-methyl group may then be incorporated by SsfM3, which is homologous to *C*-methyltransferases, to afford 2-methyl-acetoacetyl-CoA.

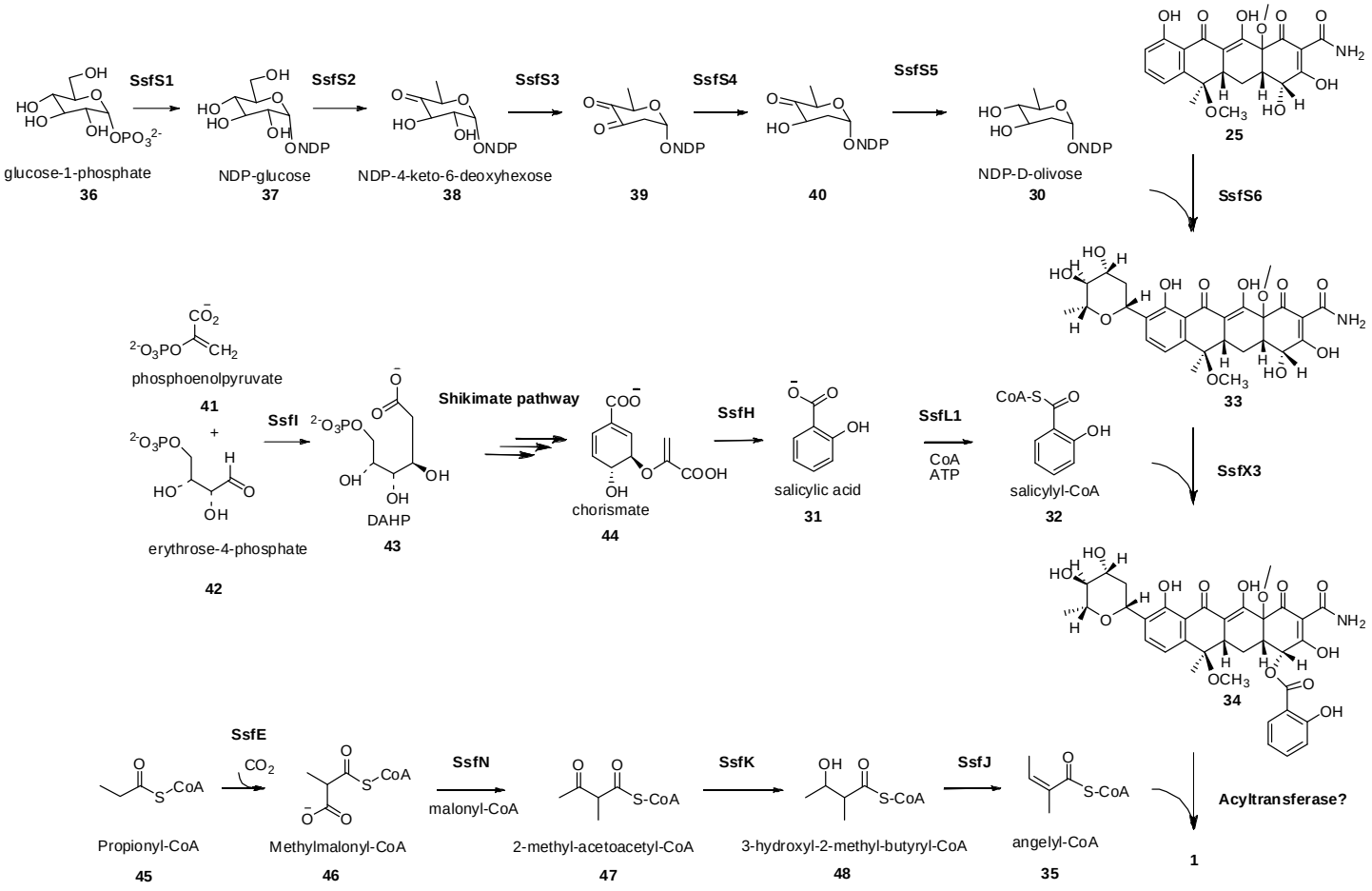

Figure 6: Proposed biosynthetic pathways for the pendants **30**, **32** and **35,** and the conversion of **25** to **1**. The order of these modifications is proposed to be glycosylation of **25** to form **33**, sialycylation of the C-4 hydroxyl of **33** to form **34**, and angelycylation of D-olivose to form **1**. The proposed intermediates **33** and **34** have been detected in the extract of *S.* sp. SF2575 culture by LCMS.

Genes involved in the biosynthesis of salicylate

The C-4 salicylate substitution in **1** is unusual not just among tetracyclines but quite rare among aromatic polyketides. One known example is that of thermorubin, in which the salicylate moiety is

thought to be incorporated as a starter unit[32]. Sequence analysis of *ssf* gene cluster revealed a likely mechanism of salicylate synthesis from the shikimate pathway (Figure 6), similar to that utilized in the biosynthesis of siderophores yersinobactin[33] and mycobactin[34, 35]. SsfH was found to be homologous to the salicylate synthase genes Irp9[33] and MtbI[34] from yersinobactin and mycobactin biosynthesis respectively, which convert chorismate, a byproduct of the shikimate pathway, to salicylate in a two step reaction that proceeds through an isochorismate intermediate. In addition, the *ssf* gene cluster contains *ssfI* that encodes a 3-deoxy-d-*arabino*-heptulosonate-7-phosphate (DAHP) synthase that condenses erythrose-4-phosphate and phosphoenolpyruvate into DAHP[36], which is the first step in the shikimate pathway. Genes encoding the remaining of the shikimate pathway that converts DAHP to chorismate are absent from the *ssf* gene cluster and are likely shared with the endogenous metabolism of *S.* sp. SF2575. The *ssfI* gene is therefore likely an extra copy dedicated for the biosynthesis of **1** and serves to direct carbon flow through the shikimate pathway when SF2575 biosynthesis is induced.

Attachment of **31** to the aglycon is proposed to occur through an ATP-dependent activation of salicylate to salicylyl-CoA **32** by a salicylyl-CoA ligase, and subsequent transfer catalyzed by an acyltransferase. SsfL1 has high sequence homology to NcsB2[37] and MdpB2[38], which catalyze formation of the CoA-ester of naphthanoic acid and 3,6-dimethylsalicylic acid in the enediyne biosynthetic pathways of neocarzinostatin and maduropeptin, respectively. To reconstitute the synthesis of salicyl-CoA, putative salicylyl-CoA ligase SsfL1 was expressed from *E. coli* BL21(DE3) as an *N*-terminal hexahistidine tag fusion protein and purified using Ni-NTA chromatography. ATP and salicylic acid were added to SsfL1 and the release of pyrophosphate (PP_i) was monitored spectrophotometrically at 340 nm using a coupled assay that oxidizes NADH. As shown in Figure 7A, PP_i release was observed only in the presence of SsfL1, ATP and salicylic acid. To demonstrate the CoA ligase properties of SsfL1, enzymatic synthesis of salicyl-CoA was performed by introducing CoA into the assay. After overnight incubation, a new compound consistent with the retention time of salicyl-CoA was purified from the reaction mixture by HPLC and mass analysis confirmed the identity ($[M+H]^+$ at m/z = 888).

Task 2. Probe the mechanism of key enzymes involved in the biosynthesis of SF2575.

Since the D-olivose biosynthetic pathway has been well established, we sought to investigate the more unclear aspects of SF2575 biosynthesis, the attachment of the unique acyl groups. There are two predicted acyltransferase enzymes in the *ssf* gene cluster, SsfX3 and SsfV. In vitro experiments have verified SsfX3, which has weak sequence homology to the GDSL lipase family[39], as the salicyl acyltransferase. To probe the potential reaction catalyzed by SsfX3, we first assayed the reverse hydrolysis reaction using different products. SF2575 was treated with purified SsfX3 overnight and the assay mixture was analyzed by LCMS. The result showed the presence the parent compound SF2575 and a second compound with mass corresponding to the loss of salicylate. A control reaction without SsfX3 did not afford this compound, as determined by selected ion monitoring. The UV of this compound showed λ_{max} of 358 nm and the loss of the shoulder at 302 nm, consistent with the hydrolysis of salicylic acid. This result was the first indication that SsfX3 may be involved in C-4 salicylation. To confirm the SsfX3-catalyzed acyltransfer reaction, an in vitro assay was performed in which SsfX3 and SsfL1 were added to **33** with salicylic acid, ATP and free CoA. The reaction was incubated for 30 minutes at room temperature and analyzed by HPLC (Figure 7B). The reaction containing all components showed nearly quantitative conversion of **33** to **34** (Figure 7B, trace iii). Omission of any of the enzyme, substrate, or cofactor component of the assay did not lead to the synthesis of **34**, as shown in Figure 7B traces iv-viii. Together, these results confirmed that SsfX3 is responsible for the conversion of **33** to **34** using **32** as the acyl donor. Since GSDL family enzymes are often found to have

broad substrate specificity, which is proposed to be the result of a flexible substrate binding pocket[39], investigating the substrate tolerance of SsfX3 may be a useful avenue for diversifying the C-4 functionality of SF2575.

Task 3. Reconstitute the biosynthesis of SF2575.

Due to the large size of the SF2575 biosynthetic gene cluster, we were unable to find a cosmid that contained the entire gene cluster. Efforts are ongoing to clone each of two cosmids that span the entire gene cluster into two different Streptomyces shuttle vectors that can be transformed simultaneously into a heterologous host, however this has not yet been successful. In place of heterologous biosynthesis, we have shifted efforts toward creating genetic knockouts to investigate the biosynthesis and create analogs of SF2575. With the successful construction of a *ssfB* knockout, we believe this strategy may be applied to many additional genes. We have begun to knockout *ssfH* in order to create a mutant that is deficient in salicylate biosynthesis which may be used to create SF2575 analogs by mutasynthesis.

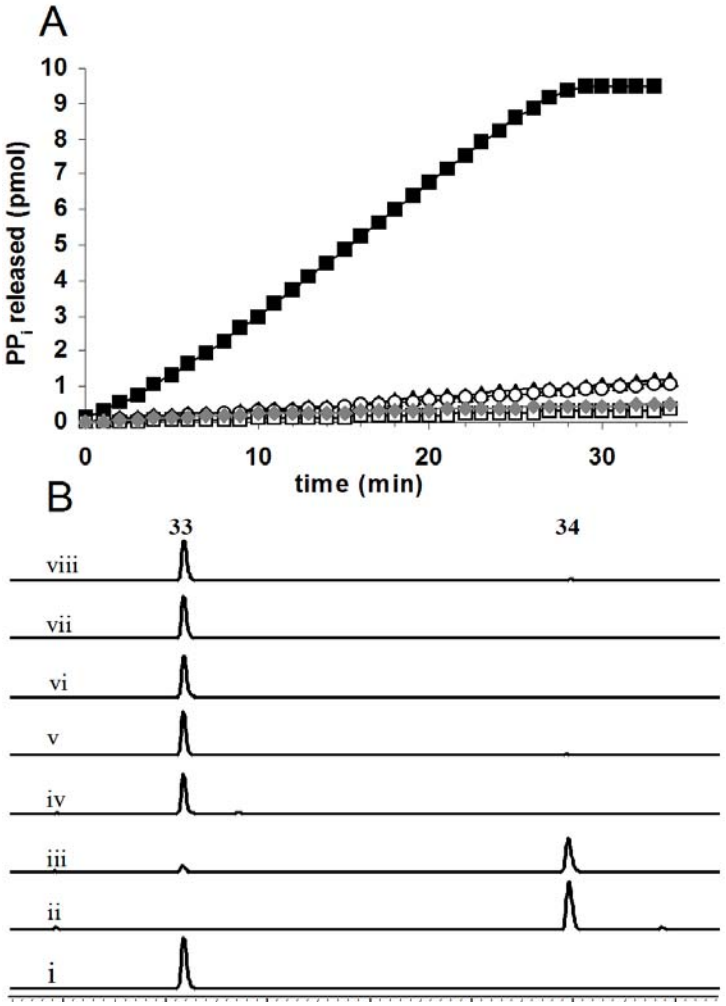

Figure 7: Reconstitution of the synthesis and transfer of salicylate to the aglycon. (**A)** Synthesis of salicylyl-AMP as indicated by the release of PP_i in the presence of ATP, salicylate and putative salicylyl-CoA ligase SsfL1. Black squares: all reaction components; black triangles: no salicylate; open circles: no ATP; open squares: no SsfL1; and grey diamonds: pyrophosphate reagent only. (**B)** The tandem actions of SsfL1 and SsfX3 transfers salicylate to the aglycon substrate **33** to yield **34.** The assays are performed in 50mM HEPES, pH 7.9 and 10 mM $MgCl_2$. i) the semisynthetic **33** standard; ii) the semisynthetic **34** standard; iii) Complete reaction containing 50 mM HEPES pH 7.9, 10 mM $MgCl_2$, 2 mM ATP, 2 mM free CoA, 2 mM **31**, 20 μM **33**, 1.5 μM SsfX3 and 15 μM SsfL1. Control reactions were performed as iii) with the following exclusions: iv) no SsfL1; v) no ATP; vi) no **31**; vii) no SsfX3; and viii) no CoA. All reactions were performed at 25°C for 30 minutes,

Task 4. Test the antitumor activities of SF2575.

Structure-activity relationship bioactivity assays

The angelate and salicylate modifications of SF2575 are unique structural features among tetracyclines and bacterial aromatic polyketides. To probe whether these moieties are the "warheads" that contribute to the antitumor activities of SF2575**,** we performed in vitro cytotoxicity assays with Nalm-6 pre-B cells using SF2575 with both acyl groups removed **33**, SF2575 with salicylic acid group hydrolyzed **34** and. While the parent compound SF2575 had a potent IC_{50} of 8.8 nM, removal of

the angelate resulted in a significantly attenuated IC_{50} of 327 nM for **34**. Further hydrolysis of the salicylate as in **33** led to an additional 15-fold decrease of potency with an IC_{50} of 5.2 μM**.** Similar trends were observed when other cell lines, including MCF-7, HeLa, and M249 were subjected to the cytotoxicity assays. Hence, both pendant groups are critically important for the bioactivity of SF2575 and are attractive targets for structural-activity-relationship studies. Tailoring enzymes identified here SsfL1 and SsfX3 responsible for the salicylate addition, and further investigation into the mechanism of addition of angelate may therefore provide useful biosynthetic approaches towards functionalizing these positions, especially considering the densely functionalized tetracycline core may be difficult to assess synthetically.

		Log(IC_{50}) Std. Error		
Cell line	Type	**1**	**34**	**33**
Nalm-6	human B cell precursor leukemia	-8.056 ± 0.0518	-6.485 ± 0.0753	-5.285 ± 0.0958
HeLa	human cervix carcinoma	-7.466 ± 0.0769	-6.196 ± 0.153	> -5.0
MCF-7	human breast adenocarcinoma	-7.619 ± 0.112	-5.068 ± 0.0813	> -5.0
M249	human melanoma	-8.134 ± 0.144	-6.788 ± 0.0921	> -5.0

Table 1: Cytotoxicity of SF2575 and derivatives

Key Research Accomplishments

- Gene cluster responsible for the biosynthesis of tetracycline family anticancer compound SF2575 has been identified and sequenced from the genome of native host *Streptomyces sp.* SF2575
- Bioinformatic analysis of genes found in the *ssf* gene cluster as well as analysis of intermediates in S. sp. SF2575 fermentation broth has been used to construct a putative pathway for the biosynthesis of SF2575
- Gene cluster has been verified through heterologous reconstitution of tetracycline producing genes through the biosynthesis of characteristic intermediate 6-methylpretetramid
- Enzymes responsible for the unique addition of salicylic acid to the A ring of SF2575 have been positively identified and characterized through in vitro analysis.
- SAR studies have been conducted by testing the bioactivity of authentic SF2575 and variants produced by removal of the angelate ligand and both the angelate and salicylate ligands. MTS assay using four different cell lines showed a marked decrease in bioactivity following the removal of the angelate and even further decrease by the removal of both acyl groups

Reportable Outcomes

- Manuscript in preparation

Lauren B. Pickens, Peng Wang, Hui Zhou, Woncheol Kim, Kenji Watanabe, Shuichi Gomi, Yi Tang "Biochemical Analysis of the Biosynthetic Pathway of an Anticancer Tetracycline SF2575".

- Presentations:

Pickens, L. B., Tang, Y., " Investigating the biosynthesis of a tetracycline analog with antitumor properties". International Conference on Biological Engineering, Jan 18-21, 2009, Santa Barbara, CA (poster)

Pickens, L. B., Tang, Y., " Investigating the biosynthesis of a tetracycline analog with antitumor properties". AICHE annual meeting, Nov 17-21, 2008, Philadelphia, PA (presentation)

- GenBank submission

SF2575 gene cluster deposited to GenBank. Accession number: GQ409537

Conclusions

We have identified and sequenced the gene cluster responsible for the biosynthesis of tetracycline family anticancer compound SF2575 from native host *S.* sp. SF2575. Using heterologous biosynthesis we have verified the tetracycline nature of the gene cluster by reconstituting the biosynthesis of key early intermediates. We have identified and characterized key tailoring enzymes responsible for the salicylic acid modification. These enzymes are unique to tetracycline biosynthetic pathways. Initial substrate specificity experiments with these enzymes indicate that they may potentially be exploited to generate improved analogs of this already potent anticancer compound, as well as the antibiotic tetracyclines.

References

1. Hatsu, M.; Sasaki, T.; Watabe, H.; Miyadoh, S.; Nagasawa, M.; Shomura, T.; Sezki, M.; Inouye, S.; Kondo, S., *J. Antibiot.* **1992,** 45, 320.
2. Horiguchi, T.; Hayashi, K.; Tsubotani, S.; Iinuma, S.; Harada, S.; Tanida, S., *J. Antibiot.* **1994,** 47, 545.
3. Kellner, U.; Sehested, M.; Jensen, P. B.; Gieseler, F.; Rudolph, P., *Lancet Oncol.* **2002,** 3, 235.
4. Pommier, Y., *Nat. Rev. Cancer* **2006,** 6, 789.
5. Zhang, W.; Ames, B. D.; Tsai, S. C.; Tang, Y., *Appl. Environ. Microbiol.* **2006,** 72, 2573.
6. Zhang, W.; Watanabe, K.; Cai, X.; Jung, M. E.; Tang, Y.; Zhan, J., *J. Am. Chem. Soc.* **2008,** 130, 6068.
7. Zhang, W.; Watanabe, K.; Wang, C. C. C.; Tang, Y., *J. Biol. Chem.* **2007,** 282, 25717.
8. Hopwood, D. A., *Chem. Rev.* **1997,** 97, 2465.
9. Wawrik, B.; Kerkhof, L.; Zylstra, G. J.; Kukor, J. J., *Appl. Environ. Microbiol.* **2005,** 71, 2232.
10. Hatsu, M.; Sasaki, T.; Gomi, S.; Kodama, Y.; Sezaki, M.; Inouye, S.; Kondo, S., *J. Antibiot.* **1992,** 45, 325.
11. Hunter, I. S., Tetracyclines. In *Microbial secondary metabolites: biosynthesis, genetics and regulation*, Martin, J. F., Ed. Research Signpost: Lucknow, India, 2002; pp 141.
12. Ichinose, K.; Bedford, D. J.; Tornus, D.; Bechthold, A.; Bibb, M. J.; Peter Revill, W.; Floss, H. G.; Hopwood, D. A., *Chem. Biol.* **1998,** 5, 647.
13. Ichinose, K.; Ozawa, M.; Itou, K.; Kunieda, K.; Ebizuka, Y., *Microbiology* **2003,** 149, 1633.
14. Menéndez, N.; Nur-e-Alam, M.; Braña, A. F.; Rohr, J.; Salas, J. A.; Méndez, C., *Chem. Biol.* **2004,** 11, 21.
15. Daum, M.; Peintner, I.; Linnenbrink, A.; Frerich, A.; Weber, M.; Paululat, T.; Bechthold, A., *ChemBioChem* **2009,** 10, 1073.
16. Prado, L.; Lombó, F.; Braña, A. F.; Méndez, C.; Rohr, J.; Salas, J. A., *Mol. Gen. Genet.* **1999,** 261, 216.
17. Brown, J. R.; Ireland, D. S., Structural Requirements for Tetracycline Activity. In *Advances in Pharmacology and Chemotherapy*, Garattini, S.; Goldin, A.; Hawking, F.; Kopin, I. J., Eds. Academic Press: New York, 1978; Vol. 15, pp 161.
18. Butler, M. J.; Friend, E. J.; Hunter, I. S.; Kaczmarek, F. S.; Sugden, D. A.; Warren, M., *Mol. Gen. Genet.* **1989,** 215, 231.
19. Nakano, T.; Miyake, K.; Endo, H.; Dairi, T.; Mizukami, T.; Katsumata, R., *Biosci. Biotechnol. Biochem.* **2004,** 68, 1345.
20. Salas, J. A.; Méndez, C., *Trends Microbiol.* **2007,** 15, 219.
21. Fernandez, E.; Weissbach, U.; Reillo, C. S.; Brana, A. F.; Mendez, C.; Rohr, J.; Salas, J. A., *J. Bacteriol.* **1998,** 180, 4929.
22. Westrich, L.; Domann, S.; Faust, B.; Bedford, D.; Hopwood, D. A.; Bechthold, A., *FEMS Microbiol. Lett.* **1999,** 170, 381.
23. Faust, B.; Hoffmeister, D.; Weitnauer, G.; Westrich, L.; Haag, S.; Schneider, P.; Decker, H.; Kunzel, E.; Rohr, J.; Bechthold, A., *Microbiology* **2000,** 146, 147.
24. Trefzer, A.; Pelzer, S.; Schimana, J.; Stockert, S.; Bihlmaier, C.; Fiedler, H. P.; Welzel, K.; Vente, A.; Bechthold, A., *Antimicrob. Agents Chemother.* **2002,** 46, 1174.
25. Bililign, T.; Hyun, C.-G.; Williams, J. S.; Czisny, A. M.; Thorson, J. S., *Chem. Biol.* **2004,** 11, 959.
26. McGaw, B. A.; Woolley, J. G., *Phytochemistry* **1979,** 18, 1647.
27. Seigler, D. S., *Plant Secondary Metabolism*. Kluwer Academic Publishers: Norwell, MA, 1998; p 759.
28. Magnuson, K.; Jackowski, S.; Rock, C. O.; Cronan, J. E., *Microbiol. Rev.* **1993,** 57, 522.
29. Tang, Y.; Lee, T. S.; Kobayashi, S.; Khosla, C., *Biochemistry* **2003,** 42, 6588.
30. Grimm, A.; Madduri, K.; Ali, A.; Hutchinson, C. R., *Gene* **1994,** 151, 1.
31. Heath, R. J.; Rock, C. O., *J. Biol. Chem.* **1996,** 271, 1833.
32. Aragozzini, F.; Craveri, R.; Maconi, E.; Ricca, G. S.; Scolastico, C., *J. Chem. Soc., Perkins Trans. 1* **1988**, 1865.
33. Kerbarh, O.; Ciulli, A.; Howard, N. I.; Abell, C., *J. Bacteriol.* **2005,** 187, 5061.
34. Zwahlen, J.; Kolappan, S.; Zhou, R.; Kisker, C.; Tonge, P. J., *Biochemistry* **2007,** 46, 954.
35. Harrison, A. J.; Yu, M.; Gårdenborg, T.; Middleditch, M.; Ramsay, R. J.; Baker, E. N.; Lott, J. S., *J. Bacteriol.* **2006,** 188, 6081.
36. Knaggs, A. R., *Nat. Prod. Rep.* **1999,** 16, 525.
37. Cooke, H. A.; Zhang, J.; Griffin, M. A.; Nonaka, K.; Van Lanen, S. G.; Shen, B.; Bruner, S. D., *J. Am. Chem. Soc.* **2007,** 129, 7728.
38. Van Lanen, S. G.; Oh, T.-j.; Liu, W.; Wendt-Pienkowski, E.; Shen, B., *J. Am. Chem. Soc.* **2007,** 129, 13082.
39. Akoh, C. C.; Lee, G.-C.; Liaw, Y.-C.; Huang, T.-H.; Shaw, J.-F., *Prog. Lipid Res.* **2004,** 43, 534.

Supporting Data

Name	Closest homolog	Putative Role	Size	Identity/ Similarity	Accession number
SsfA	OxyA	Ketosynthase	45 kDa	76% / 86%	CAA80985
SsfB	OxyB	Chain length factor	43 kDa	64% / 77%	AAZ78326
SsfC	OxyC	Acyl carrier protein	9.5 kDa	54% / 73%	AAZ78327
SsfD	OxyD	Amidotransferase	68 kDa	68% / 79%	AAZ78328
SsfE	ChlJ	Carboxytransferase	58 kDa	81% / 89%	AAZ77684
SsfF	Npun_F4742	Ketoreductase	34 kDa	55% / 70%	ACC83094
SsfG	ChlB3	Ketosynthase III	36 kDa	37% / 56%	DQ116941
SsfH	MbtI	Salicylate synthase	48 kDa	42% / 57%	Q7D785
SsfI	PlmI	DAHP-7-phosphate synthase	45 kDa	56% / 67%	AY354515
SsfJ	SAV_2786	Enoyl CoA hydratase/isomerase	28 kDa	61% / 75%	BAC70497
SsfK	Athe_1156	3-Ketoacyl-(ACP) reductase	26 kDa	38% / 59%	ACM60257
SsfL1	SdgA	Salicyl-AMP ligase	58kDa	75% / 84%	BAC78380
SsfL2	OxyH	Acyl-CoA ligase	56 kDa	51% / 67%	DQ143963
SsfM1	OxyT	O-methyl transferase	36 kDa	48% / 62%	DQ143963
SsfM2	DauK	O-methyl transferase	35 kDa	38% / 53%	AAB16938
SsfM3	MtmMII	Methyl transferase	37 kDa	38% / 55%	CAK50790
SsfM4	OxyF	C-6 methyl transferase	37 kDa	51% / 68%	AAZ78330
SsfN	PpzT	Ketosynthase III	31 kDa	46% / 63%	CAX48662
SsfO1	OxyS	Oxygenase	55 kDa	42% / 54%	AAZ78342
SsfO2	OxyL	Oxygenase	59 kDa	57% / 70%	AAZ78335
SsfP	CmmQ	Dehydrogenase	29 kDa	49%/54%	YP_118212
SsfQ	Lct52	SAM synthetase	43 kDa	84% / 90%	ABX71135
SsfR	Krad_0275	MFS transporter	58 kDa	33% / 51%	ABS01765
SsfS1	StrD	dTDP glucose synthase	39 kDa	72% / 83%	BAG22759
SsfS2	CalS3	Glucose-4,6-dehydratase	36 kDa	69% / 80%	AAM94770
SsfS3	Sim20	Hexose-2,3-dehydratase	53 kDa	58% / 72%	AF322256
SsfS4	ChlC4	Hexose-3-ketoreductase	33 kDa	49% / 59%	AAZ77681
SsfS5	NanG4	Hexose-4-ketoreductase	37 kDa	42% / 53%	AAP42863
SsfS6	HedJ	Glycosyltransferase	41 kDa	40% / 55%	AAP85354
SsfT1	SnorA	Regulation	31 kDa	44% / 61%	CAA12016
SsfT2	SAML0351	Regulation (TetR family)	23 kDa	48% / 62%	CAJ89338
SsfU	OxyJ	C-9* ketoreductase	27 kDa	76% / 86%	AAZ78333
SsfV	ZhuC	Acyl transferase	34 kDa	54% / 64%	AAG30190
SsfX1	ChlI	Unknown function	8.5 kDa	50% / 60%	AAZ77683
SsfX2	RemJ	Carbohydrate kinase	36 kDa	64% / 72%	CAE51179
SsfX3	AviX9	Acyltransferase	40 kDa	48% / 61%	AAK83171
SsfY1	OxyK	Aromatase/Cyclase	34 kDa	57% / 67%	AAZ78334
SsfY2	NcnE	Second ring cyclase	33 kDa	56% / 68%	AAD20271
SsfY3	CmmQ	Cyclase/aromatase	31 kDa	29% / 42%	CAE17552
SsfY4	OxyI	Cyclase/aromatase	13 kDa	66% / 76%	AAZ78332
Orf-1	Hypothetical protein SGR_5601	Unknown	12 kDa	84%/94%	BAG22430
Orf-2	SSEG_08767	FeS assembly protein	17 kDa	89%/96%	EDY55253
Orf-3	SSEG_03816	Cysteine desulferase	46 kDa	90%/94%	EDY65440
Orf1	Hypothetical conserved protein	Unknown	66 kDa	60%/69%	EEP13034
Orf2	Hypothetical protein SCO3803	Unknown	14 kDa	78%/88%	NP_627922
Orf3	SSEG_08979	AraC Transcriptional Regulator	35 kDa	82%/90%	EDY56197

Table A1 Putatively assigned functions of enzymes encoded in the *ssf* biosynthetic gene cluster